DOCUMENT RESUME

ED 175 905 TM 009 261

AUTHOR Threadgill, Judith
TITLE The Interaction of Learner Aptitude with Types of Questions Accompanying a Written Lesson on Logical Implications.
PUB DATE [Apr 79]
NOTE 'p.: Paper presented at the Annual Meeting of the American Educational Research Association (63rd, San Francisco, California, April 8-12, 1979)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Academic Aptitude; Foreign Countries; Grade 10; *Learning Processes; *Mathematical Logic; Performance Factors; *Questioning Techniques; Secondary School Mathematics; Senior High Schools; *Verbal Ability; Vocabulary
IDENTIFIERS *Aptitude Treatment Interaction; Canada

ABSTRACT

Since research on the effect of questions adjunct to written materials has usually ignored learner aptitudes, this study was designed to explore the interaction of a verbal aptitude measure (vocabulary ability) with two types of questions following a prose selection on logical implications. Two hundred fifty-two tenth-grade students were randomly assigned to one of four treatments: text with applied questions, text with verbatim questions, text with no questions, and control text. Regression analysis yielded an ordinal interaction, suggesting that applied questions were more effective than no questions for students with low aptitude scores, while not differentially affecting students with high aptitude scores. In addition, a disordinal interaction was identified which suggested that students with low aptitude scores learned more from verbatim questions, while students with high aptitude scores learned more with no questions. (Author/RD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED175905

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

The Interaction of Learner Aptitude With Types of Questions Accompanying a Written Lesson on Logical Implications

Judith Threadgill

University of Calgary

Calgary, Alberta

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Judith Threadgill

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

TM009 261

Abstract

Mathemagenic research on the effect of questions adjunct to written materials has usually ignored learner aptitudes. This study was designed to explore the interaction of a verbal aptitude measure with two types of questions following a prose selection on logical implications. Two hundred fifty-two tenth-grade students were randomly assigned to one of four treatments: text with applied questions, text with verbatim questions, text with no questions, and control text. Regression analysis yielded an ordinal interaction ($p = .029$) suggesting that applied questions were more effective than no questions for students with low aptitude scores, while not differentially affecting students with high aptitude scores. In addition, a disordinal interaction ($p = .058$) was identified which suggested that students with low aptitude scores learned more from verbatim questions while students with high aptitude scores learned more with no questions.

The Interaction of Learner Aptitude With Types of Questions Accompanying a Written Lesson on Logical Implications

The purpose of this study was to explore the differential learning effects of two types of questions following mathematical prose materials. Learning from meaningful prose materials through the insertion of adjunct questions has been the subject of numerous recent investigations (for reviews, see Faw & Waller, 1976; Wilson & Koran, 1976). Interest in the effect of questions on learning was prompted by Rothkopf's (1965) introduction of the notion of mathemagenic behaviors, that is, those attending behaviors which give birth to learning. Factors which might be said to produce mathemagenic behaviors in the student involve instructional materials, settings, and directions (Arter & Clinton, 1974). However, only studies using the technique of inserting questions to achieve particular instructional objectives have grown directly out of Rothkopf's theory (Faw & Waller, 1976).

Several critical variables have been identified which influence learning from prose materials. Questions occurring after material to which they relate tend to increase both incidental and intentional learning (Frase, 1970; Rothkopf, 1966). The same effects have been found when retention measures are used (Sanders, 1973). In addition, the type of questior posed also determines performance levels. More learning occurs with general questions than with comparative or specific questions (Frase, 1968), and with higher order questions than lower

order questions (Shavelson, Berliner, Ravitch & Loeding, 1974, Watts & Anderson, 1971). Finally, a small number of questions presented adjunct to a lesson can be less effective than no questions at all (Holliday, Note 1).

Other variables remain relatively unexplored or, at most, their appearance in the literature raises more issues than it clarifies. A specific example of such a variable is the effect of differing ability levels on questioning treatments, with the concomitant aptitude-treatment interactions (ATI) which can occur. Very few such interactions have been found and inconsistencies occur in some of these. For example, Shavelson et al. (1974) were able to find an interaction where students of low ability performed better with higher order questions following the reading of prose material, while high ability students obtained better scores when assigned the text without questions. However, no such interaction existed in a replication study which extended the treatment from a few prose passages to an entire textbook (Shavelson, Berliner, Loeding, Porteus & Stanton, Note 2). In a somewhat inconclusive study, Sanders (1973) found postquestions better for high ability students and prequestions better for low ability students on an immediate retention test, but the opposite disordinal interaction appeared on a delayed test. Using the pacing of questions as a variable, Koran & Koran (1975) found frequent questions more profitable for lower ability students while spaced questions or no questions were more effective for high ability students on a measure of incidental learning, but there

was no such interaction on a measure of relevant retention. It should be noted that questions used in this study required learners to remember highly specific types of factual information, and that the interaction effect was likely due to this type of question. Finally, in a study more relevant to the present one in terms of treatment content, Kuehls (1976) found that below average students profited more from a matrix algebra text passage with inserted questions, whereas the questions had little effect on performance of high ability students. Unfortunately, the type of question used in the study (high order or low order) was not specified. No ability measure was used; students were either in an honors course or nonhonors course.

These studies lead to the apparent conclusion that procedures designed to promote mathemagenic behaviors can be counterproductive for some learners (Cronbach & Snow, 1977). Such a conclusion is unwarranted at this time, since so few ATI studies have been conducted in the area of questions adjunct to a text. It is of considerable importance to understand more clearly the effect of inserted questions on learning by people of different ability levels before these ideas can be implemented into actual classroom settings (Faw & Waller, 1976).

Except for the Kuehl's study (1976) all of the investigations reported above concentrated on the instructional effect of the questions, with little regard given to the actual content of the prose selections as related to a particular discipline. Yet the processes involved in reading written materials can be highly dependent upon the discipline

from which the materials are taken. This is especially true in mathematics, where symbolism, page structure, extended vocabulary and many other factors call for reading skills not necessary to other disciplines (Kane, Byrne & Hater, 1973). The ability to read textual materials in mathematics takes on increasing importance as classrooms become less teacher dominated (Forbes, 1970). It has been suggested that students who can become involved in self-teaching acquire much more independence in learning (Catterson, 1975), and that the degree of students' skill to read mathematics carefully and analytically is highly related to their mathematical achievement (Aiken, 1972).

Accordingly, the present study was designed to examine possible interactions between a general aptitude measure (vocabulary) and two question types adjunct to mathematical textual materials. Questions followed a prose selection on logical implications. The two types of questions selected for this study were based on descriptions of test items suggested by Anderson (1972). The applied questions asked for identification of new instances of the logical patterns introduced in the text, whereas the verbatim questions were taken in literal word-by-word form from the text. Ordinal interactions were hypothesized where applied questions would prove more effective than verbatim questions, which would in turn be more effective than no questions for students scoring low on the aptitude measure. On the other hand, students scoring high on the aptitude measure would be little affected by question type or lack of questions.

Method

Sample

The 252 students participating in the investigation were enrolled in nine classes of Mathematics 10, from three public high schools in Calgary, Alberta, Canada. Mathematics 10 is a standard Alberta course for grade ten students, and covers traditional algebra and geometry content.

Materials

The main instructional treatment consisted of a lesson on logical implications, adapted from Harnadek (1969). Four logical inference patterns were presented: 1) Law of detachment, 2) Denying the antecedent, 3) Affirming the consequent, and 4) Contrapositive inference. Each pattern was first presented as a syllogism in sentence form with the corresponding symbolic form, followed by an explanation of why the pattern was valid or invalid. Two more syllogisms followed in sentence form, using nonsense phrases, where the symbolic pattern was applied to the phrases to show why the argument was valid or invalid.

Twelve applied questions were prepared, three from each pattern, and randomly distributed on the question sheet. For each question two sentences were given, for example, "If algebra is fun, then cows are purple. Cows are purple." Students were asked to identify p and q. (In this case, p is the phrase "algebra is fun", q is the phrase "Cows are purple".) They were then asked to identify the second sentence as p, q, not p, or not q, identify the pattern used, and state a logical

conclusion if the pattern was a valid one. Twelve verbatim questions were also prepared. In each instance, students were given two lines of a syllogism from the text. They were asked to locate the syllogism in the text, supply the missing sentence, and note the pattern used and its validity. Again, three syllogisms were taken from each of the four patterns. The control group material was also adapted from Harnadek (1969), and consisted of prose selections on circular reasoning and special pleading, with questions asking students to identify instances of the two types of reasoning. These questions had the simple function of keeping control students occupied during treatment sessions, and were not considered adjunct questions for purposes of analyses.

The posttest contained 20 multiple-choice items. Each item consisted of the first two sentences of a syllogism and four possible answers. The first three response alternatives were always possible conclusions, and the fourth was always the statement "No conclusion can be made". Each of the four patterns was represented by five items. Within each group of five items, one was a straightforward application of the pattern, two were nonsense syllogisms, and two contained negations in the first sentence. The 20 items were randomly distributed on the test.

The instrument used as the aptitude measure was the 36-item vocabulary test (V-2) from the Kit of Reference Tests for Cognitive Factors (French, Ekstrom & Price, 1963). Vocabulary tests in general yield high positive correlations with arithmetic reasoning and general intelligence (Aiken, 1972), and the V-2 test in particular is widely used as a measure

of general ability (French, et al., 1965; Shavelson et al., 1974).

Procedure

Within each of the nine sections of Mathematics 10, students were randomly assigned to one of the four treatments, text with applied questions (AQ), text with verbatim questions (VQ), text with no questions (NQ) or control (C). Each treatment was printed on a different color paper, and students were told that each color represented a different lesson on logical thinking, some with questions and some with no questions required to understand the material. Explanation of the different colors of paper was given so that students would not look at nearby students' work and differentiate the materials. Students were not allowed to have any other materials on their desks, and were instructed to spend the entire alloted time on the lessons. Students asking for assistance on questions were given individual help on the first one or two questions, and students who had no questions and finished early were instructed to review. No other feedback was given. Except for a few students who completed only nine or ten of the twelve applied questions, all students were able to complete the lesson within the 45 minutes allowed. The vocabulary (V-2) test (8 minutes) was administered immediately preceding the lessons, and the posttest (15 minutes) was administered immediately following the treatment lessons. Although all students received the same posttest, the posttest distributed to each student was printed on paper the same color as the treatment. Students were therefore under the impression that each test matched the particular

lesson each had completed.

Results

Descriptive statistics for the V-2 test (aptitude measure) and the posttest (achievement measure) are reported in Table 1.

Place Table 1 about here

One-way analysis of variance was used to test instructional treatment main effects, resulting in a significant F ratio, F (3,248) = 20.112, $p < .01$. The Scheffe method ($p < .05$) was then used to compare pairs of means (Kerlinger & Pedhazur, 1973). The control treatment mean proved to be significantly lower than the other three treatment means. Since the function of the control group was to provide a baseline for evaluating the effects of the other treatments, it was subsequently dropped from further analyses. Of the remaining treatment groups, the applied question treatment mean proved to be significantly higher than the verbatim question and the no question treatment means, but the verbatim question and no question treatment means were not significantly different.

Furthermore, posttest scores were regressed on aptitude scores for all three possible pairwise combinations of the remaining three treatment groups, thus evaluating the ATI hypotheses. In the first analysis, no significant interaction was found for the applied question and verbatim question treatments, F (1,124) = 1.75, p = .188. It was therefore

assumed that the slopes of the regression lines for these two treatment groups were not significantly different.

When posttest scores were regressed on aptitude scores for the applied question and no question treatments, a significant F value was found for the variance increment due to interaction, F (1,124) = 4.89, p = .029. Because this significant F implies that regression lines for the two treatment groups have significantly different slopes, regression equations for each treatment group were calculated and the lines plotted (see Figure 1).

Insert Figure 1 about here

Since in this instance the point of intersection (25.1, 19.9) was barely within the range of scores, the interaction was interpreted as ordinal.

In the third analysis, posttest scores were regressed on the aptitude scores for the verbatim question and no question treatments. Although the F value in this instance did not attain a high significance level, F (1,124) = 3.66, p = .058, analysis was continued due to the relatively low probability value of F. The point of intersection (16.2, 14.7) of these two regression lines was within both ranges of scores, and therefore the interaction was interpreted as disordinal (see Figure 2).

Insert Figure 2 about here

Discussion

The most effective treatment for the entire sample was the text followed by application-type questions, as predicted. This result is not surprising, since the applied questions demanded deeper analysis of textual material than was required of the other main treatments, and therefore facilitated comprehension of the logical patterns involved. Verbatim questions or lack of questions did not require students to semantically encode, that is, to process the meaning of the text (Anderson, 1972) to the same degree as applied questions. One could perhaps interpret this result as an example of the "law of least effort"; people can be counted on to engage only in the processing demanded by the task (Anderson, 1970).

The ordinal interaction found between the applied question treatment and the no question treatment sheds further light on the above result. Although it is true that applied questions proved more effective than no questions, the interaction indicates that this difference was far more prominent for students at the lower end of the aptitude scale. Ironically though, classroom teachers more often choose to give lower order questions rather than higher order questions to students of low ability, perhaps acting on the premise that since these students can more successfully answer lower order questions, they are more apt to learn more from these questions. However, the present result implies that educators might be better advised to ask higher order questions of low ability students, and be less concerned about the type of questions

given to high ability students, who are often better able to function regardless of the instructional questioning technique.

These results are in partial agreement with those of the Shavelson et al. (1974) study. Their recommendations following a disordinal interaction were that students with low vocabulary scores be assigned to a text with higher orde questions, while students with high vocabulary scores might be assigned to a text without questions. However, no such interaction was found in the replication study (Shavelson et al., Note 2), nor was the interaction disordinal in the present study.

Interaction results comparing the applied question treatment with the verbatim question treatment did not result in the hypothesized interaction. It is possible that the lack of significant interaction is due to a Type II error, since it was not possible to obtain a large enough sample to reduce the Type II error probability (Cronbach and Snow, 1977). Although the interaction was not significant, the direction was again towards low ability students performing better with applied questions than verbatim questions, and the effect of question type being minimized as ability increased.

Providing a rationale for the disordinal, rather than the hypothesized ordinal interaction found between the verbatim question treatment and the no question treatment is more difficult. Apparently, verbatim questions facilitated learning for low ability students by providing a method (albeit an incomplete one) of organizing information and focusing attention on relevant information. However, such questions proved to be

dysfunctional for high ability students. Perhaps these students felt inappropriately secure about their comprehension of the textual material due to the ease with which they were able to answer the verbatim questions. On the other hand, the no question treatment placed students in a more uncertain situation, with the concomitant demand to semantically encode the information to a degree of personal satisfaction that the material was understood, a demand which only the high ability students were able to accomplish satisfactorily. Certainly, encoding information without the aid of questions is a demanding task, and learning depends on general ability to a greater degree when active intellectual work is required of the learner (Cronbach & Snow, 1977).

Only one aptitude measure was examined in this aptitude-treatment interaction study. Although performance on this particular aptitude measure (vocabulary) is indicative of verbal and general ability, which correlate positively with mathematical ability, it would be profitable to examine possible interactions between these or similar treatments and mathematical aptitude measures. Mathematical reasoning ability and problem solving ability would seem to be good indicators of success on treatments involving logical thinking, and it would be of interest to evaluate whether similar interactions occur using tests measuring these aptitude constructs.

It would also be of interest to examine treatments more closely aligned with traditional textbook approaches and topics in which symbolism and mathematical applications are more common than in the treatments

of this study. Textbook explanations of mathematical concepts and principles are often not assigned to be read by students, since many teachers prefer to provide oral explanations and assign only exercises or problems (NACOME, 1975). Yet the accompanying information in the text is presumably subjected to at least as much rigorous planning and editing as most teacher explanations, and the two can complement one another to provide a clearer exposition of the content of a lesson. Through additional mathemagenic research on the effectiveness of questions adjunct to mathematical text, it may be possible to provide teachers and textbook writers with information needed to make mathematical prose more meaningful.

Reference Notes

1. Holliday, W. G. Differential learning effects of varied numbers of instructive questions. Paper presented at the meeting of the National Association for Research in Science Teaching, Cincinnati, March 1977.

2. Shavelson, R. J., Berliner, D. C., Loeding, D., Porteus, A. W., & Stanton, G. C. Adjunct questions, mathemagenics and mathemathanics. Paper presented at the meeting of the American Psychological Association, New Orleans, September 1974.

References

Aiken, L. R. Language factors in learning mathematics. Review of Educational Research, 1972, 42, 359-385.

Anderson, R. C. Control of student mediating processes during verbal learning and instruction. Review of Educational Research, 1970, 40, 349-369.

Anderson, R. C. How to construct achievement tests to assess comprehension. Review of Educational Research, 1972, 42, 145-170.

Arter, J. A., & Clinton, L. Time and error consequences of irrelevant data and question placement in arithmetic word problems II: fourth graders. Journal of Educational Research, 1974, 68, 28-31.

Catterson, J. H. Techniques for improving comprehension in mathematics. In G. G. Duffy (Ed.), Reading in the Middle School. Newark, Delaware: International Reading Association, 1974.

Cronbach, L. J., & Snow, R. E. Aptitude and Instructional Methods. New York: Irvington Publishers, 1977.

Faw, H. W., & Waller, T. G. Mathemagenic behaviors and efficiency in learning from prose materials: Review, critique, and recommendations. Review of Educational Research, 1976, 46, 691-720.

Forbes, J. E. Textbooks and supplementary materials. In M. F. Rosskopf (Ed.), The Teaching of Secondary School Mathematics: Thirty-Third Yearbook. Reston, Virginia: National Council of Teachers of Mathematics, 1970.

Frase, L. T. Effect of question location, pacing, and mode upon retention of prose material. Journal of Educational Psychology, 1968, 59, 244-249.

Frase, L. T. Boundary conditions for mathemagenic behaviors. Review of Educational Research, 1970, 40, 337-347.

French, J. W., Ekstrom, R. B., & Price, L. A. Kit of Reference Tests for Cognitive Factors. Princeton, N.J.: Educational Testing Service, 1963.

Harnadek, A. Mathematical Reasoning. Birmingham, Michigan: Midwest Publications, 1969.

Kane, R. B., Byrne, M. A., & Hater, M. A. Helping Children Read Mathematics. New York: American Book Company, 1974.

Kerlinger, F. N., & Pedhazur, E. J. Multiple Regression in Behavioral Research. New York: Holt, Rinehart and Winston, 1973.

Koran, M. L., & Koran, J. J. Interaction of learner aptitudes with question pacing in learning from prose. Journal of Educational Psychology, 1975, 67, 76-82.

Kuehls, E. A. Effect of interspersed questions on learning from mathematical text. Journal for Research in Mathematics Education, 1976, 7, 172-175.

National Advisory Committee on Mathematical Education. Overview and Analysis of School Mathematics Grades K-12. Reston, Virginia: National Council of Teachers of Mathematics, 1975.

Rothkopf, E. Z. Some theoretical and experimental approaches to problems in written instruction. In J. D. Krumbolz (Ed.), Learning and the Educational Process. Chicago: Rand McNally, 1965.

Rothkopf, E. Z. Learning from written instructive materials: An exploration of the control of inspection behavior by test-like events. American Educational Research Journal, 1966, 3, 241-249.

Sanders, J. R. Retention effects of adjunct questions in written and aural discourse. Journal of Educational Psychology, 1973, 65, 181-186.

Shavelson, R. J., Berliner, D. C., Ravitch, M. M., & Loeding, D. Effects of position and type of question on learning from prose material. Journal of Educational Psychology, 1974, 66, 40-48.

Watts, G. H., & Anderson, R. C. Effects of three types of inserted questions on learning from prose. Journal of Educational Psychology, 1971, 62, 387-394.

Wilson, J. J., & Koran, J. J. Review of research on mathemagenic behavior: Implications for teaching and learning science. Science Education, 1976, 60, 391-400.

Table 1

Means, Standard Deviations and Reliabilities

of Vocabulary Test and Posttest

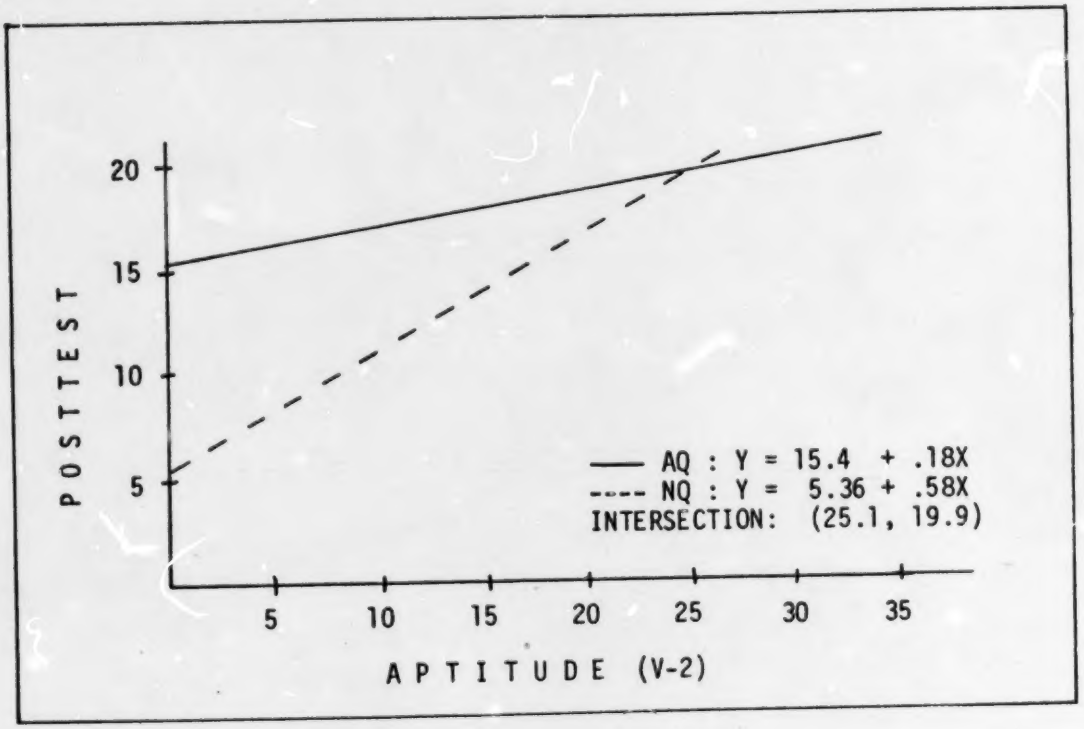
Test	Reliability	Treatment Group			
		AQ	VQ	NQ	C
		N=63	N=63	N=63	N=63
V-2	.76	18.45(4.57)	18.66(4.76)	19.12(5.83)	19.12(5.13)
Posttest	.86	18.73(5.16)	15.60(4.89)	16.38(5.35)	12.33(2.87)

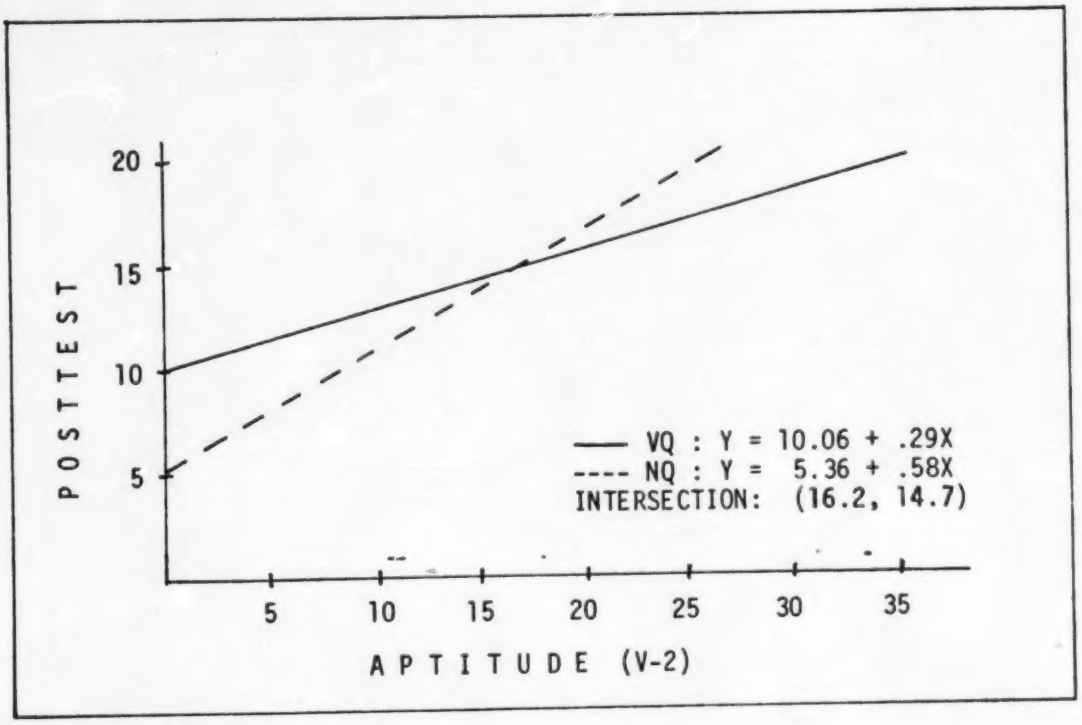
Note. AQ = Applied Questions, VQ = Verbatim Questions, NQ = No Questions, C = Control.

Figure Captions

Figure 1. Interaction of aptitude with applied question and no question treatments.

Figure 2. Interaction of aptitude with verbatim question and no question treatments.


20
15
10
5
POSTTEST
AQ : Y = 15.4 + .18X
NQ : Y = 5.36 + .58X
INTERSECTION: (25.1, 19.9)
5
10
15
20
25
30
35
APTITUDE (V-2)


20
15
10
5
POSTTEST
VQ : Y = 10.06 + .29X
NQ : Y = 5.36 + .58X
INTERSECTION: (16.2, 14.7)
5
10
15
20
25
30
35
APTITUDE (V-2)